

CS474 Project Midterm

Date: 2025-03-27 13:11

#HOMEWORK

#CS474

Linear Models & SVM applied to Health Dataset

Content

- INTRODUCTION
 - RELATED WORKS
 - METHODS
 - PRELIMINARY RESULTS
 - FUTURE PLAN
 - REFERENCES
-

Introduction

The purpose of this project is to both increase our knowledge of the implementations of linear models, support vector machine, and gradient descent, while applying them to a real world dataset.

Our choice of dataset is a heart disease csv found on Kaggle, which is a cumulation of three different health organization's data. Features included age, gender, resting heart rate, resting blood pressure, as well as numerous other nomial features, and a binary target feature indicating if that patient had heart disease.

Our focus is to implement each model using different subsets of features; linear regression model to predict a scalar value such as blood pressure using age and heart rate, perceptron model to determine if a patient has heart disease, logisitic regression to determine the probability of a patient having heart disease, and a support vector machine to compare the accuracy against the perceptron. Because we can't assume that our linear regression RSS will be invertible, gradient descent will also be implemented.

Our predicted outcomes are that the linear regression model may under perform due to the many other factors influencing biometrics, as well as not being able to fully utilize the nomial features, which will also effect the logisitic regression & perceptron for the same reasons (the underlying driver for each linear model is regression.) Due to this the support vector machine

may actually be a semi-useful predictor for heart disease, especially when compared to the perceptron.

Related Works

So far we have just been using the class notes and textbook; we will search for research papers for the final write up.

Methods

Data Processing

```
def load_features(filepath, features=["age", "sex", "resting bp s"], target=
["target"], remove=False):
```

```
def split_data(features, labels, reduce=False):
```

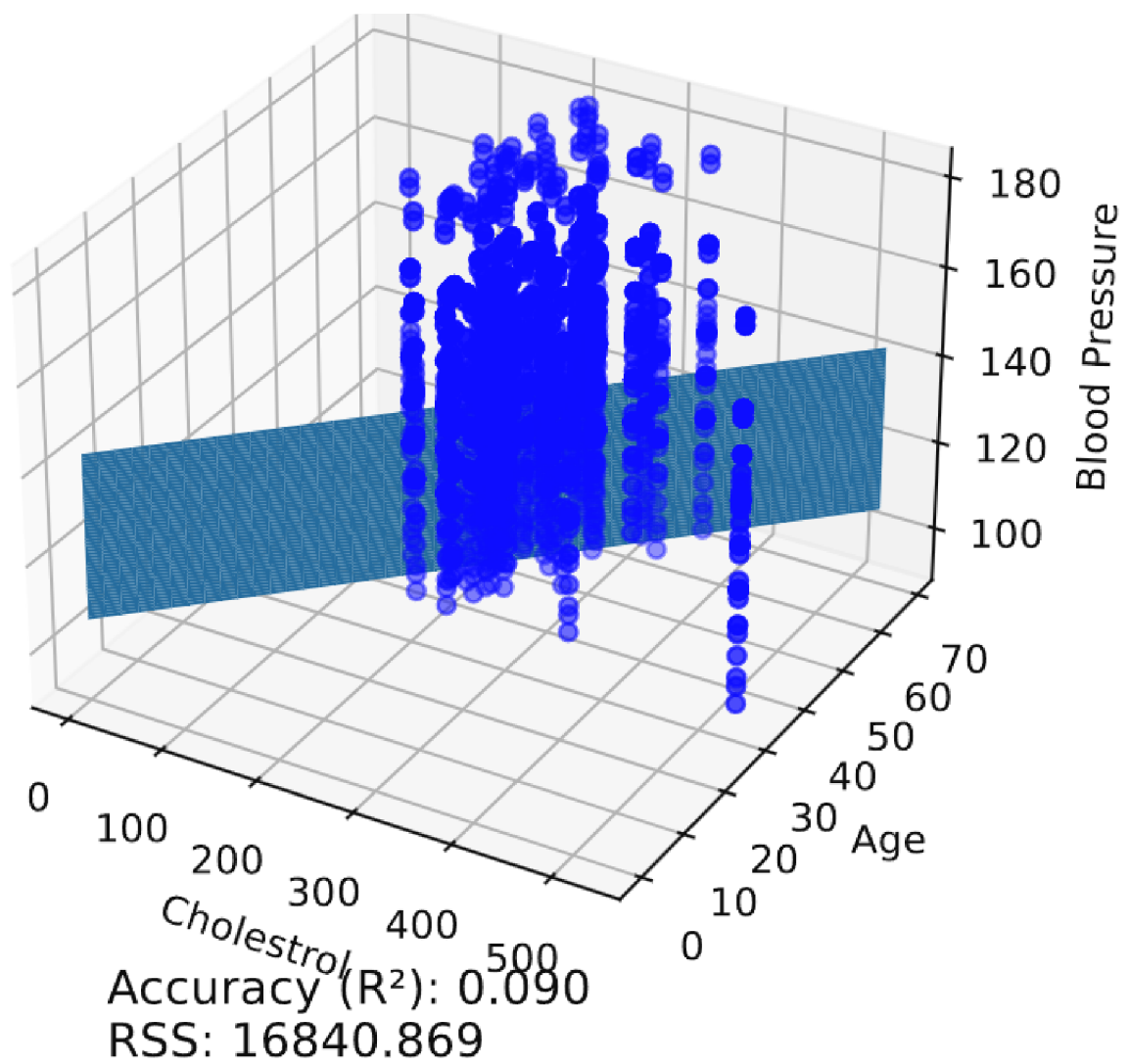
Linear Regression

```
def fit(self, X, y)
def accuracy(self, X, y)
def predict(self, X)
```

Linear Perceptron

```
def fit(self, X, y)
def accuracy(self, X, y)
def predict(self, X)
```

Preliminary Results



Future Plan

References

Yaser S. Abu-Mostafa, Malik Magdon-Ismail, Hsuan-Tien Lin, "Learning From Data"
AMLbook.com
